

1. 2. 3. 4. 5. 6.

(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MPearch<sub>jp</sub> protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Apr 1 11:02:14 1996; MasPar time 5.08 Seconds

Tabular output not generated.

Title: >US-08-347-748-2

Perfect Score:

Sequence: 1 MEETELLVWMLLTARTL.....PTSP LINTSYTHSQNLSEQE 353

Scoring table: PAM 150

## Gap 11

Searched: 62355 seqs, 7230759 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: a-geneseq20

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8 9:part9 10:part10 11:part11 12:part12

Statistics: Mean 34.391; Variance 182.424; scale 0.189

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query Match	Length	ID	Description	Pred. No
No.			DB			
1	146	5.9	193	R11893	Erythropoietin analog	2_41e-02
2	144	5.8	193	R11891	Erythropoietin analog	3_26e-02
3	141	5.7	188	P60599	Clone Lambda HRPOTfL6	5_15e-02
4	141	5.7	188	P81195	Erythropoietin encode	5_15e-02
5	141	5.7	193	R65499	Human prepro-erythropoietin	5_15e-02
6	141	5.7	193	P60597	Clone Lambda HRPOTfL3	5_15e-02
7	141	5.7	193	R11892	Erythropoietin analog	5_15e-02
8	141	5.7	193	P50300	Human erythropoietin	5_15e-02
9	141	5.7	193	P70256	Sequence of human erythropoietin	5_15e-02
10	141	5.7	330	R23076	Epo:IL-3 short, recombinant	5_15e-02
11	141	5.7	349	R23079	Epo:IL-3 Flex, recombinant	5_15e-02
12	140	5.6	193	P60598	Open reading frame cDNA	5_99e-02
13	140	5.6	193	R11858	Modified human erythropoietin	5_99e-02

14	134	5.4	193	3	P50344	EPO encoded by clone	4.17e-011
15	127	5.1	192	3	P50301	Monkey erythropoietin	1.47e-011
16	128	5.1	193	2	R11859	Erythropoietin analog	3.60e-011
17	126	5.1	503	6	R34127	Anexin XI type I ffo	4.83e-011
18	121	4.9	166	12	R63647	Mouse erythropoietin.	1.00e+000
19	119	4.8	302	4	R23596	Recombinant hematopoi.	1.34e+000
20	119	4.7	321	4	R23075	IL-3:Epo short, recom	1.34e+000
21	116	4.8	314	4	R23597	Recombinant hematopoi	2.07e+000
22	116	4.7	340	4	R23078	IL-3:Epo Flex, recomb	2.07e+000
23	115	4.6	77	1	P81199	Erythropoietin encode	2.39e+000
24	114	4.6	168	12	R63560	Pig erythropoietin.	2.76e+000
25	115	4.6	3144	11	R58777	Protein encoded by Hu	2.39e+000
26	113	4.5	165	12	R63645	Monkey erythropoietin	3.18e+000
27	113	4.5	166	12	R63644	Human erythropoietin.	3.18e+000
28	113	4.5	166	2	P70398	Sequence of human ery	3.18e+000
29	113	4.5	166	4	R23593	Recombinant hematopoi	3.18e+000
30	113	4.5	167	3	P50299	Human recombinant ery	3.18e+000
31	113	4.5	167	3	P50298	Human recombinant ery	3.67e+000
32	112	4.5	167	2	R07665	Cysteine-added variin	3.67e+000
33	113	4.5	303	4	R23598	Recombinant hematopoi	3.18e+000
34	111	4.5	322	4	R23599	Recombinant hematopoi	3.18e+000
35	111	4.5	334	3	P83194	Sequence of a bioadhe	4.23e+000
36	109	4.4	165	12	R63648	Rat erythropoietin.	5.61e+000
37	109	4.4	505	6	R56646	Anexin XI type I ffo	5.61e+000
38	109	4.4	610	9	R56664	Mutant platelet glyco	5.61e+000
39	109	4.4	610	9	R51116	Platelet glycoprotein	5.61e+000
40	106	4.3	165	12	R63646	Monkey erythropoietin	8.55e+000
41	107	4.3	166	12	R62603	Cat erythropoietin.	7.43e+000
42	107	4.3	439	5	R28150	Sugar beet chitinase	7.43e+000
43	106	4.3	732	3	R47043	Mammalian son of seve	8.55e+000
44	105	4.2	749	8	R14311	Human milk bile-salt	9.83e+000
45	105	4.2	744	8	R45189	BSTL/CTL.	9.83e+000

## ALIGNMENTS

RESULT	1
ID	R11893 standard; Protein; 193 AA.
AC	R11893;
DE	22-JUL-1991 (first entry)
DT	Erythropoietin analogue [Pro]124,[Thr]125.
KM	Human erythropoietin; EPO; Isoform; analogue; haematocrit;
KM	sialic acids; glycosylation.
OS	Homo sapiens.
PN	W09105867-A.
PN	EP-428267-A.
PD	02-MAY-1991.
PE	09-OCT-1990; U05758.
PR	13-OCT-1989; US-421444.
PA	(AMGE-) AMGEN INC.
PI	Strickland TW, Byrne TE, Elliott SG;
DR	WPI; 91-148745/20.
DR	WPI; 91-150265/21.
PT	Recombinant erythropoietin iso-forms and purificn. - increase
PT	haematocrit levels in mammals and contg. specific number of
PT	sialic acids
CC	Claim 33; Page 45; 60pp; English
CC	The analogue was constructed to add an O-glycosylation site
CC	at Thr125 (amino acid 154 in this sequences) and was produced by
CC	expression of DNA obtained by site-directed mutagenesis of DNA
CC	encoding EPO. The analogue has a higher sialic acid content than
CC	human EPO and has increased biological activity.
CC	It can be used to treat mammals to cause bone marrow cells to
CC	increase prodm. of reticulocytes and red blood





PD 02-MAY-1991.  
 PF 09-OCT-1990; U05758.  
 PF 13-OCT-1989; U542144.  
 PA (AMGE-) AMGEN INC.  
 PT Strickland TW, Byrne TE, Elliott SG;  
 DR WPI; 91-148745/20.  
 DR WPI; 91-150265/21.  
 PT Recombinant erythropoietin iso-forms and purification. - increase  
 PT haematocrit levels in mammals and contg. specific number of  
 -PT sialic acids  
 PS Claim 33; Page 45; 60pp; English  
 CC The analogue was constructed to add an O-glycosylation site  
 CC at Thr125 (amino acid 154 in this sequence) and was produced by  
 CC expression of DNA obtained by site-directed mutagenesis of DNA  
 CC encoding Epo. The analogue has a higher sialic acid content than  
 CC human Epo and has increased biological activity.  
 CC It can be used to treat mammals to cause bone marrow cells to  
 CC increase prodn. of reticulocytes and red blood  
 CC thereby increasing haematocrit levels.  
 CC See also R11859, R11891-93.  
 SQ Sequence 193 AA;

Query Match	5.7%;	Score 141;	DB 2;	Length 193;
Best Local Similarity	27.6%;	Pred. No. 5.15e-02;		
Matches	27;	Conservative	27;	Mismatches 41; Indels 3; Gaps 3;

Dbb  
12 LLLSLSLPLGLVIGAPRLICSDSVLERYLLEAKAENITGCAEHCNINITYPDT 71  
||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||  
7 LLVVMLLTARLT-ISSPARACDLRVLSKLRDSHVLHSRISQCEVHNPILPVVLLPAV 65

```

Db      72 kvnfyawh-rmevgqavevngqllalseavlrgall 108
      :: || :: | :: ::|| |:: :: |
Oy      66 DESLGEKTKQME-ETKQADILCAVTTLLLEGVMARQQL 102

```

RESULT 8  
ID P50300 standard; protein; 193 AA.

DT 01-JAN-1980 (first entry)  
DE Human erythropoietin encoded by positive clone (phage lambda-hE1)  
DE isolated from human fetal liver gene bank.

KW	disorder; ss; phage lambda-hel; gene bank.
OS	Homo sapiens.
PN	W08502610-A.

PF 11-DEC-1984; U02021.  
PR 13-DEC-1983; US-561024.  
PR 21-FEB-1984; US-582185.

PR 30-NOV-1984; US-675298.  
PA (KIRI-) KIRIN-AMGEN INC.  
DR WPT; 85-159229/26.

PT New polypeptide having properties of erythropoietin - is prepd.  
PT by cultivation of transformed eucaryotic or procaryotic host  
PS Disclosure; Page 43; 113pp; English.

CC lambda-hE1 is essential for red blood cell formation and is used  
CC for the diagnosis and treatment of blood disorders such as anaemia  
CC large amounts of EPO may be obtained using recombinant DNA

CC techniques in contrast to small amounts obtained from plasma  
CC and urine. This sequence is expressed in *E. coli*. See also  
CC N50345-6, N50348-50 and P50298-99, P50301.

**SQ** Sequence 193 AA;

Query Match 5.7%; Score 141; DB 3; length 193;  
Best Local Similarity 27.6%; Pred. No. 5.15e-02;  
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

**D**b    12 llslslslplqivlgaprrllcdstveryllleakeaenittgcaehslmenitpydt 71  
      || : | | ::| : | | | : : : : | : | :  
**Q**y    7 LLVVMLLLTARLT-LSSPAPACOLRVSLKLTDSHVLSRLSQCPENHPLPTPVLLPAV 65

```

Db 72 kmfiyawk-rmevvgqavevnglallseavlryqall 108
    :: ||::| |::::||::|::|::|
Qy 66 DFLSGEMKTYME-ETKQDILCAVTLLEGVMARQQL 102

```

RESULT	9
ID	P70256 standard; protein; 193 AA.

DT	19-FEB-1991 (first entry)
DE	Sequence of human erythropoietin (EPO).
KW	Renal anaemia therapy; hormone.

	Key	Location/Qualifiers
FH	Peptide	1..27
FT		
FT	/label=SIGNA	

FT	Region	81.97
FT	/note="Fragment that probe N70361 is based on"	
PN	EP-232034-A.	

PF 19-JAN-1987; 300399.  
PR 23-JAN-1986; JP-012868.  
PA (SUMO) SUMITOMO CHEM IND KK

PI Yanagi H, Ogawa I, Okamoto M, Hozumi T, Soga A, Yoshima T,  
PI Teutsumi M.  
DR WPI, 87-223006/32.

PT Human erythropoietin prodn. - by culturing human cells, esp.  
PT Namalwa cells, transformed with DNA encoding human erythropoietin  
PS Disclosure; Fig 1; 22pp; English.

from the erythropoietin-producing human hepatoma cell Hp-1. The cDNA library was screened using the probes given in N70361 and N70362. A plasmid (named as p58-A20) was isolated. The nucleotide sequence of

Query Match	Score	DB 2	length
Sequence	193	AA;	

Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

7 LLMVMLLTARLT-LSSPAPACDLRVLSKLEDSHVLSRSLSGCEVNP LPTPLPAV 65

66 DFLGEMKQME-ETKQDILGAVTLLLEGVMARGOL 102

RESULT	10
ID	R23076 standard; Protein; 330 AA
AC	R23076;

DT	20-OCT-1992 (first entry)
DE	Epo:IL-3 short, recombinant haematopoietic molecule.
KM	Early MDF; Late MDF; haematopoiesis; Epo; IL-3; growth factor.
OS	Homo sapiens.
EH	Key Location/Qualifiers
FT	Peptide 1..27
FT	/label= sig_peptide
FT	protein 28..330
FT	/label= mat_protein
PN	M09206116-A.
PD	16-APR-1992.
PF	26-SEP-1991; U07053.
PR	28-SEP-1990; US-589958.
PA	(ORTH ) ORTHO PHARM CORP.
PI	Rosen JI;
DR	WPI; 92-150819/18.
DR	N-PDB; Q24282.
PT	Recombinant haematopoietic molecules useful in treating
PT	anaemia(s) - comprise IL-3 or GM-CSF, EPO, G-CSF, IL-5 or M-CSF
PS	and has early and later myeloid differentiation activity
PT	Disclosure; Page 44; 82pp; English.
CC	The amino acid sequence given is an Epo:IL-3 hybrid growth factor
CC	derived from a construction formed by ligating the native Epo signal
CC	sequence and various synthetic oligonucleotides corresponding to Epo
CC	and IL-3 gene sequences. This hybrid growth factor is a haematopoietic
CC	molecule which contains at least a portion of an early MDF and at least
CC	a portion of a late MDF covalently linked. This compound can be used
CC	to promote haematopoiesis in a patient.
CC	The bonding of the early and late factors allows a very high conc. of
CC	late MDF at the surface of a cell which the early MDF is bound. It
CC	also allows the early MDF to act more specifically to stimulate only
CC	the desired lineage, thus reducing undesirable effects. These
CC	compounds are useful for treating anaemias of various origins eg. renal
CC	failure and AIDS. It is easier to produce and administer one
CC	recombinant molecule rather than two separate molecules.
CC	Sequence 330 AA;

	Query Matchn	5.7%	Score 141;	DB 4;	Length 330;
	Best Local Similarity	27.6%;	Pred. No.	5.1Se-02;	
Matches	27;	Conservative	27;	Mismatches	41; Indels 3; Gaps 3;
D8	12 Lllellslplqpvlgaprilcdsrlyrlleakeaeenlttgcacelcslnmetivtpct 71   :: :- - :::    ::          ::  : - :-  :				
Oy	7 LIIVMLLLIARLT-LSSPAPRACDILRVLSKLRIDSHVISHRLSQCEYHPLPTPVILLPAV 65				
D9	72 kmfyawk-rmevvggagvwmgglallsaelirgaall 108 ::   ::   :- :: ::    - :-  :: :				
Oy	66 DFLSGEMKTOME-ETKAODILGAATVLLECGWAMARGOL 102				
RESULT	11				
ID	R23079 standard; Protein; 349 AA.				
AC	R23079;				
DE	20-OCT-1992 (first entry)				
DT	Epo:IL-3 Flex, recombinant hematopoietic molecule.				
KM	Early MDF; late MDF; haematopoesis; Epo; IL-3; linker; growth factor.				
OS	Homo sapiens.				
FH	Key Location/Qualifiers				
FT	Peptide 1..27				
FT	/label= sig_peptide				
FT	Protein 28...349				
FT	/label= mat_protein				
NN	MO9206116-A.				

PD 16-APR-1992.  
 PF 26-SEP-1991; U07053.  
 PR 28-SEP-1990; US-589958.  
 PA (ORTH ) ORTHO PHARM CORP.  
 P1 Rosen JI;  
 DR WP1; 92-150819/18.  
 DR N-PSDB; 024285.  
 PT Recombinant haematopoietic molecules useful in treating  
 PT anaemia(s) - comprise IL-3 or GM-CSF, Epo, G-CSF, IL-5 or M-CSF  
 PT and has early and later myeloid differentiation activity  
 PS Disclosure; Page 51; 82pp; English.  
 CC The amino acid sequence given is an Epo:IL-3 hybrid growth factor  
 CC derived from a construction formed by ligating the native Epo signal  
 CC sequence and various synthetic oligonucleotides corresponding to Epo  
 CC and IL-3 gene sequences. This molecule is comparable to the sequence  
 CC given in R23076 and contains a flexible linker molecule. This hybrid  
 CC growth factor is a haematopoietic molecule which contains at least a  
 CC portion of an early MDF and at least a portion of a late MDF  
 CC covalently linked. This compound can be used to promote hematopoiesis  
 CC in a patient.  
 CC The bonding of the early and late factors allows a very high conc. of  
 CC late MDF at the surface of a cell which the early MDF is bound. It  
 CC also allows the early MDF to act more specifically to stimulate only  
 CC the desired lineage, thus reducing undesirable effects. These  
 CC compounds are useful for treating anaemias of various origins eg. renal  
 CC failure and AIDS. It is easier to produce and administer one  
 CC recombinant molecule rather than two separate molecules.  
 SO Sequence 349 AA;

	Query Match	5.7k; Score 141; DB 4; Length 349;
	Best Local Similarity	27.6%; Pred. No. 5.13e-02;
	Matches	27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;
Db	12	lplglpvlgpprlcdsvlyeryllekeaeenltgcacnslneinivcpt 71   ::     ::         ::  ::  ::  ::  ::  ::  ::  ::  ::   Qy 7 LLLVMLLTLRLT-ISSPAPRACDLRVLSKLRDSHVLRSLSCPEVRHPLRPVLLAV 65
Db	72	kvnfyawk-rimevgaqevawvggllalseavlrqall 108 ::  :     ::  ::   ::  ::   Qy 66 DFLSGEKTKTOME-ETKKADITLGAVTLLIEGVMAARQL 102
RESULT	12	
ID	P60598	standard; protein, 193 AA.
AC	P60598:	
DT	01-JAN-1980	(first entry)
DE	Open reading frame coding for the erythropoietin tryptic fragment of lambda HEP01.	
DE	Erythropoietin; lambda HEP01; recombinant plasmid vector; anaemia; mammal cell culture; fTJ; CHO; Chinese hamster ovary; ss. Homo sapiens. MN08603520-A.	
PD	19-JUN-1986.	
PF	03-DEC-1985; U02405.	
PR	04-DEC-1984; US-677813.	
PR	03-JAN-1985; US-688622.	
PR	22-JAN-1985; US-693258.	
PA	(GENE-) GENETICS INST INC.	
PA	(FRIT/) FRITSCH E.	
P1	Fritsch E, Hewick RM, Jacobs K;	
DR	MPI: 86-169459/26.	
DR	N-PEDBj. N60518.	
T7	Prodn. of human cDNA clone expressing erythropoietin - for mass prodn. of erythropoietin. useful for treating anaemia	



**SQ** Sequence 192 AA;

Best Local Similarity	25.3%	Pred. No. 4.1/e-01;
Matches	25; Conservative	30; Mismatches 41; Indels 3; Gaps 3,

66 DESIGENKTOME-ETKAQDILGAVTLLLEGVMARGOLG 103

Job time : 44 secs.

(a) (b) (c) (d) (e)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MParch\_pp protein - protein database search, using Smith-Waterman algorithm

```
Run on:      Mon Apr 1 11:01:07 1996;  MasPar time 11.78 Seconds
            257 EEO will be 11.78 seconds
```

Tabular output not generated.

Title: >US-08-347-748-2

Perfect Score: 2493

Sequence: I MELTELLLVMLLTARITL.....PTSPPLNTSYTHSQNSQEG 353

Scoring table: PAM 150  
Gap 11

Searched: 82306 seqs, 25270970 residues

Post-processing: Minimum Match 0%

Database: plr46

```
1:ann1 2:ann2 3:ann3 4:unann1 5:unann2 6:unann3 7:unann4
8:unann5 9:unann6 10:unann7 11:unann8 12:unrev1
13:unrev2
```

Statistics: Mean 46.752; Variance 133.030; scale 0.351

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description	Pred. No
1	2493	100.0	353	13	S48740	thrombopoietin - hum	0.00e+00
2	2493	100.0	353	13	S45331	c-MET Ligand - human	0.00e+00
3	1748	70.1	356	13	S45330	thrombopoietin - mou	3.38e-23
4	1711	68.6	326	13	JC4125	thrombopoietin - rat	1.08e-22
5	1143	45.8	286	10	A55330	megakaryocyte growth	1.06e-14
6	155	6.2	192	5	A24902	erythropoietin precu	3.61e-05
7	154	6.2	192	13	S28148	erythropoietin - rat	4.65e-05
8	141	5.7	193	2	Z0VHU	erythropoietin precu	1.19e-03
9	139	5.6	503	2	LURB11	amexin XI - rabbit	1.94e-03
10	132	5.3	485	8	A33647	sulfated surface gly	1.05e-02
11	127	5.1	192	5	J00173	erythropoietin precu	3.41e-02
12	127	5.1	193	5	A24901	erythropoietin precu	3.41e-02

13	126	5.1	503	2	LUBO11	annexin XI form A -	4.30e-02
14	121	4.9	522456			hydroxyproline-rich	1.36e-01
15	121	4.9	505	10	A53152	annexin XI - human	1.36e-01
16	116	4.7	228	12	S55504	extensin-like protel	4.21e-01
17	116	4.7	817	9	S51342	verprolin - yeast (S	4.21e-01
18	115	4.6	297	12	S23737	proline-rich protein	5.26e-01
19	115	4.6	421	13	S29599	Acrocin (EC 3.4.21.1	5.26e-01
20	115	4.6	441	12	S44554	hypothetical protein	5.26e-01
21	115	4.6	441	9	S31861	hypothetical protein	5.26e-01
22	115	4.6	1874	1	JO1533	RNA-directed RNA pol	5.26e-01
23	115	4.6	344	10	A46068	Huntingtin's disease	5.26e-01
24	112	4.5	267	12	S08314	cell wall glycoprote	1.02e+00
25	112	4.5	328	9	JO0955	hydroxyproline-rich	1.02e+00
26	112	4.5	1232	10	S40756	hypothetical protein	1.02e+00
27	109	4.4	268	6	S06666	protein TPX-VI3 - Th	1.95e+00
28	110	4.4	473	8	S50755	hypothetical protein	1.57e+00
29	109	4.4	474	6	S15921	protein TPX-VI3 - Th	1.95e+00
30	109	4.4	505	2	S23447	annexin XI form B -	1.95e+00
31	109	4.4	626	3	NBHUYA	platelet glycoprotei	1.95e+00
32	106	4.3	153	9	S46998	hypothetical protein	3.71e+00
33	107	4.3	283	9	S13383	hydroxyproline-rich	3.00e+00
34	106	4.3	353	11	S31166	transcription factor	3.71e+00
35	106	4.3	353	11	C54258	hepatocyte nuclear f	3.71e+00
36	106	4.3	377	10	A48018	mucin, M62=low molec	3.71e+00
37	107	4.3	434	7	B55253	ferredoxin reductase	3.00e+00
38	106	4.3	434	12	S11967	nodule-specific (hyd	3.71e+00
39	107	4.3	439	8	S45025	chitinase (EC 3.2.1.	3.00e+00
40	107	4.3	464	11	A41655	spliceosome-associat	3.00e+00
41	106	4.3	593	13	S49525	glycoprotein C (homo	3.71e+00
42	106	4.3	1206	13	S24407	formin isoform IV -	3.71e+00
43	106	4.3	1336	11	S21361	hypothetical protein	3.71e+00
44	106	4.3	1336	11	S23716	hypothetical protein	3.71e+00
45	106	4.3	1468	13	S11515	formin - mouse	3.71e+00

## ALIGNMENTS

RESULT	1
ENTRY	S68740
TITLE	#type complete thrombopoietin - human
ORGANISM	#formal name Homo sapiens #common name man
DATE	01-Aug-1995 #sequence_revision 01-Aug-1995 #text_change 01-Aug-1995

## REFERENCE

**#authors**  
 Sohma, Y.; Akahori, H.; Seki, N.; Hori, T.; Ogami, K.; Kato, T.; Shimada, Y.; Kawamura, K.; Miyazaki, H.

## # Journal

**##title** Molecular cloning and chromosomal localization of the human thrombopoietin gene.

**#access**

```
##status      preliminary
##residues    1-353 ##label SOH
```

## SUMMARY

Query Match	100.0%;	Score 2493;	DB 13;	Length 353;
Best Local Similarity	100.0%;	Prod No 0	00a+00;	

## Matches

Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

## Db 1 F

Db 1 melte|||vml|||tar|||t|ssappacdlrv|sk||rdshv|hsr|sqcpevh|ptpv 60

1 N  
Ov

Qy 1 ME TEL LLV M LLT AR TL TSS P APAC D LRV LSK L RDS HV LHS RL S QCP EVHP L P TPV 600

## Db 61 1

Db 61 llpavdflgewtkmeetkadi lgavtlllegvmaarqqlgptclssllgqlsgvr l 120



```
|||||
Qy 61 LRPADVS LG EKQTQMEETKAD IIGAVT LLECGVMAARQGLPTCTSS LLAQSLSQVRL 120
Db 121 l l g a l g s l l t c t q l p p g q r t t a h k d n a i f l e f g h l l r q k v r f l m l v g g s c l v r i a p p t 180
|||
Qy 121 LIGALQSL LCTQ LPPQGR TTAHKDPNAIFLSFOHLIRGRKVRFLMVLVGSSTLCVRARAPPTT 180
Db 181 a v p s r t s l v t l n e l p n r t s g l l e t n f t a s a r t g s g l m k o o g f r a k i p e l l n o t s r s l 240
|||
Qy 181 A V P S R T S L V T L N E L P N R T S G L L E T N F T A S A R T G S G L M K O O G F R A K I P E L L N O T S R S L 240
Db 241 d d i p p y l n r t h e l n g t r q l f g p p e r t l q a p d i s s e t d g e l p p n l q p g y s p t h p p 300
|||
Qy 241 D D I P E Y L N R H E I L N G T R G L F P G P S R R T L C A P D I S S G T D G S L P N L Q P Y S P S P T H P P 300
Db 301 t e g y t l f p l p t l p t p v v q l h p l l p d p s a p t p t p s p l m n t s y t h s o n l s o e g 353
|||
Qy 301 T E G Y T L F P L P T L P T P V V Q L H P L L P D P S A P T P T P S P L M N T S Y T H S O N L S O E G 353

RESULT 2
ENTRY 2
TITLE 545331 #type complete
ORGANISM c-MPL ligand - human
#formal_name Homo sapiens #common_name man
DATE 10-Dec-1994 #sequence_revision 10-Dec-1994 #text_change
10-Dec-1994

ACCESSIONS 545331
REFERENCE 545331
#authors de Sauvage, F.J.; Haes, P.E.; Spencer, S.D.; Malloy, B.E.;
Gurney, A.L.; Spencer, S.A.; Darbonne, W.C.; Henzel, W.J.;
Wong, S.C.; Kuang, M.J.; Olee, K.J.; Hultgren, B.; Solberg
Jr., L.A.; Goedeel, D.V.; Eaton, D.L.
#journal Nature (1994) 369:533-538
#title Stimulation of megakaryocytopoiesis and thromopoiesis by the
c-Mpl ligand.
#accession 545331
#status preliminary
#residues 1-353 ##label DEA
SUMMARY #length 353 #molecular-weight 37822 #checksum 1597

Query Match 100.0%; Score 2493; DB 13; Length 353;
Best Local Similarity 100.0%; Pred. No. 0.00e+00;
Matches 353; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Db 301 t e g y t l f p l p t l p t p v v q l h p l l p d p s a p t p t p s p l m n t s y t h s o n l s o e g 353
|||
Qy 301 T E G Y T L F P L P T L P T P V V Q L H P L L P D P S A P T P T P S P L M N T S Y T H S O N L S O E G 353

RESULT 3
ENTRY 3
TITLE 545330 #type complete
ORGANISM thrombopoietin - mouse
#formal_name Mus musculus #common_name house mouse
DATE 10-Dec-1994 #sequence_revision 10-Dec-1994 #text_change
10-Dec-1994

ACCESSIONS 545330
REFERENCE 545330
#authors Lok, S.; Kaushansky, K.; Holly, R.D.; Kuiper, J.L.;
Lofgren-Day, C.E.; Oort, P.J.; Grant, F.J.; Heipel, M.D.;
Burkhead, S.K.; Kramer, J.M.; Bell, L.A.; Sprecher, C.A.;
Blumberg, H.; Johnson, R.; Prunkard, D.; Ching, A.F.T.;
Mathewes, S.L.; Bailey, M.C.; Forstrom, J.M.; Buddle, M.M.;
Osborn, S.G.; Evans, S.J.; Sheppard, P.O.; Pressnell, S.R.;
O'Hara, P.J.; Hagen, F.S.; Roth, G.J.; Foster, D.C.
#journal Nature (1994) 369:565-568
#title Cloning and expression of murine thrombopoietin cDNA and
stimulation of platelet production in vivo.
#accession 545330
#status preliminary
#residues 1-356 ##label LOK
SUMMARY #length 356 #molecular-weight 37835 #checksum 9983

Query Match 70.1%; Score 1748; DB 13; Length 356;
Best Local Similarity 71.6%; Pred. No. 3.38e-23;
Matches 255; Conservative 47; Mismatches 49; Indels 5; Gaps 4;
```

```
Db 1 m e l t e l l a m l l a v a r t l s e p p a c d r l v s k l l r d a h v l h e r l s e q e v h p l p t p v 60
|||
Qy 1 M E L T E L L A M L L T A R L T L S P A P A C D R A V L S K L R D S H V L S R L S Q C E V H P L P T P V 60
Db 61 l l p a v d f s l g e w k t c g e e k a d i l g a v t l l e g v m a a r g q l g p t c l s s l l a q s l s q v r l 120
|||
Qy 61 L L P A V D F S L G E K T Q M E E T K A D I I G A V T L L E C V M A A R Q G L P T C L S S L L A Q S L S Q V R L 120
Db 121 l l g a l g s l l t c t q l p p g q r t t a h k d n a i f l e f g h l l r q k v r f l m l v g g s c l v r i a p p t 180
|||
Qy 121 L I G A L Q S L L C T Q L P P Q G R T T A H K D P N A I F L S F O H L I R G K V R F L M V L V G S S T L C V R A R A P P T T 180
Db 181 a v p s r t s l v t l n e l p n r t s g l l e t n f t a s a r t g s g l m k o o g f r a k i p e l l n o t s r s l 240
|||
Qy 181 A V P S R T S L V T L N E L P N R T S G L L E T N F T A S A R T G S G L M K O O G F R A K I P E L L N O T S R S L 240
Db 241 d d i p p y l n r t h e l n g t r q l f g p p e r t l q a p d i s s e t d g e l p p n l q p g y s p t h p p 300
|||
Qy 241 D D I P E Y L N R H E I L N G T R G L F P G P S R R T L C A P D I S S G T D G S L P N L Q P Y S P S P T H P P 300
Db 301 t e g y t l f p l p t l p t p v v q l h p l l p d p s a p t p t p s p l m n t s y t h s o n l s o e g 353
|||
Qy 301 T E G Y T L F P L P T L P T P V V Q L H P L L P D P S A P T P T P S P L M N T S Y T H S O N L S O E G 353

RESULT 4
ENTRY 4
TITLE 3C4125 #type complete
ORGANISM thrombopoietin - rat
#formal_name Rattus norvegicus #common_name Norway rat
DATE 27-Aug-1995 #sequence_revision 27-Aug-1995 #text_change
27-Aug-1995

ACCESSIONS 3C4125
```

[illegible]

```

#map_position 3q26.3
KEYWORDS      alternative splicing; cytokine
SUMMARY       #length 286 #molecular-weight 31544 #checksum 6126

Query Match      45.8%; Score 1143; DB 10; Length 286;
Best Local Similarity 94.2%; Pred. No. 1.06e-142;
Matches 163; Conservative 2; Mismatches 6; Indels 2; Gaps 2;

Db 1 melteallvmllltarlllseappacdlrvslkllrtdahvlnherlscqcvnplptpv 60
|||||
Qy 1 MELTEALLVVMMLLTARLTLSPPAPACDLRVLSKLLRDSHVLSRLSQCPDVNR LPPTV 60

Db 61 llpavdfejgwkfimeetkcaqdligavtlllegvmaargqlptcslsllsqleqgyr1 120
|||||
Qy 61 LLPAVDFSLGEKKTQMEETKADODLIGAVTLLLEGVMAARGQLPTCLSLSQLSGGVRL 120

Db 121 llgalgsllgtcldppqgrtrtahnkdnafistfghllrpkd-fwi-vgdklhcl 171
|||||
Qy 121 LLGALGSLGTCLDPPGRTTAAHKDPNAIFLSFHLLGKGVRFLLMVGSGTLCV 173

RESULT 6
ENTRY      A24902 #type complete
TITLE      erythropoietin precursor - mouse
ORGANISM   Mus musculus #common_name house mouse
DATE       25-Oct-1987 #sequence_revision 25-Oct-1987 #text_change 18-Jun-1993

ACCESSIONS A24902
REFERENCE   A24902
#authors    Shoemaker, C.B.; Mitscock, L.D.
#journal     Mol. Cell. Biol. (1986) 6:849-858
#title       Murine erythropoietin gene: cloning, expression, and human
              gene homology.
#cross-references MIMID:87039105
#accession    A24902
##molecule_type DNA
##residues    1-192 ##label SHO
#note         the authors translated the codon TTA for residue 12 as
              Phe, TTA for residue 43 as Phe and ATG for residue 145
              as Ile

GENETICS
#introns     5/1; 52/3; 81/3; 141/3
CLASSIFICATION
KEYWORDS     erythropoiesis; glycoprotein; hormone
SUMMARY      #length 192 #molecular-weight 21339 #checksum 6155

Query Match      6.2%; Score 155; DB 5; Length 192;
Best Local Similarity 26.8%; Pred. No. 3.61e-05;
Matches 34; Conservative 35; Mismatches 54; Indels 4; Gaps 4;

Db 10 llllellllplglpvlcaprllcdsrvlerylleakeaenvtmgcaegprlseinltvad 69
||||:|:|:| 1 1:|:| 11 11: : : : : 1-1 1: : : :
Qy 6 LLVVMMLLTARLTLSPPAPACDLRVLSKLLRDSHVLSRLSQCPDVNR LPPTVLLPA 64

Db 70 tkymfyawk-rmeveqaealevwqglslseallgagallanesqpcetqlhldkaleql 128
::|:|:| 1 1:|:| 1:|:| 1:|:| 1:|:| 1:|:| 1:|:| 1:|:| 1:|:| 1:|:|
Qy 65 VDSLSGEKKTQME-ETKADODLIGAVTLLLEGVMAARGQLPTCLSLSQLSGGVRL-L 122

Db 129 raltail 135
::| 111
Qy 123 GALTSL 129

RESULT 7

```

ENTRY S28148 #type complete  
TITLE erythropoietin - rat  
ORGANISM #formal name Rattus norvegicus #common name Norway rat  
DATE 22-Nov-1993 #sequence\_revision 22-Nov-1993 #text\_change 22-Nov-1993

ACCESSIONS S28148  
REFERENCE S28148  
#authors Nagao, M.; Suga, H.; Okano, M.; Masuda, S.; Narita, H.; Ikura, K.; Sasaki, R.  
#journal Biochim. Biophys. Acta (1992) 1171:99-102  
#title Nucleotide sequence of rat erythropoietin.  
#accession S28148  
#status preliminary  
#residues 1-192 ##label NAG

SUMMARY #length 192 #molecular-weight 21286 #checksum 4587

Query Match 6.2%; Score 154; DB 13; length 192;  
Best Local Similarity 25.8%; Pred. No. 4.65e-05;  
Matches 34; Conservative 35; Mismatches 58; Indels 5; Gaps 5;

Db 10 LILLELLEPLGLPGLCAPRLLCDSEVLYLLEAKAENITGCAEHLNITPTD 69  
||| : ||| : | : | : || ||| : : : : | : | : : | :  
Qy 6 LLVVMLLTARLT-LSSPAPACDLRLVLSKLIRDSVHLRSQCPVHP LPTVLLPA 64

Db 70 tkrvfwkrmkveeqavevwqglalleaillqagqlqanagqpsslqlhickaisglr 129  
:: || : | : | : : ||| : : : || : : : || : : : | :  
Qy 65 VDFSLGEMKTQMEETKADILLGAVTLLLEGVMAARG-QLGRT-CLSSLLQSLSGVRL 121

Db 130 sltstllrvlqag 141  
| : | : ||| : |  
Qy 122 -LGAAGSLAGTG 132

RESULT 8  
ENTRY 8 #type complete  
TITLE erythropoietin precursor - human  
ORGANISM #formal name Homo sapiens #common name man  
DATE 27-Nov-1985 #sequence\_revision 27-Nov-1985 #text\_change 22-Apr-1995

ACCESSIONS A01855; A24744; A25384; A22210  
REFERENCE A01855  
#authors Jacobs, K.; Shoemaker, C.; Rudersdorf, R.; Neill, S.D.; Kaufman, R.J.; Mufson, A.; Seehra, J.; Jones, S.S.; Hewick, R.; Fritsch, E.F.; Kawakita, M.; Shimizu, T.; Miyake, T.  
#journal Nature (1985) 313:806-810  
#title Isolation and characterization of genomic and cDNA clones of human erythropoietin.  
#cross-references M01D:85137899  
#accession A01855  
#molecule\_type mRNA; DNA  
#residues 1-193 ##label JAC  
##cross-references GB:X02157; GB:X02158

REFERENCE A24744  
#authors Lin, F.K.; Suggs, S.; Lin, C.H.; Browne, J.K.; Smalling, R.; Egrie, J.C.; Chen, K.K.; Fox, G.M.; Martin, F.; Stabinsky, Z.; Badrawi, S.M.; Lai, P.H.; Goldwasser, E.  
#journal Proc. Natl. Acad. Sci. U.S.A. (1985) 82:7580-7584  
#title Cloning and expression of the human erythropoietin gene.  
#accession A24744  
#molecule\_type DNA  
#residues 1-193 ##label LIN  
##cross-references GB:M1319

REFERENCE A25384

#authors Lai, P.H.; Everett, R.; Wang, F.F.; Arakawa, T.; Goldwasser, E.  
#journal J. Biol. Chem. (1986) 261:3116-3121  
#title Structural characterization of human erythropoietin.  
#cross-references M01D:86140080  
#accession A25384  
#molecule\_type protein  
#residues 28-86, 'Q', 87-193 ##label LAI  
#experimental\_source urine  
#note forms without the carboxyl-terminal residue and the four carboxyl-terminal residues were observed

REFERENCE A22210  
#authors Yanagawa, S.; Hirade, K.; Ohnata, H.; Sasaki, R.; Chiba, H.; Ueda, M.; Goto, M.  
#journal J. Biol. Chem. (1984) 259:2707-2710  
#title Isolation of human erythropoietin with monoclonal antibodies.  
#cross-references M01D:84135751  
#accession A22210  
#molecule\_type protein  
#residues 28-29, 'X', 31-33, 'L', 35-50, 'X', 52-53, 'D', 55, 'G', 57

COMMENT Erythropoietin is produced by kidney or liver of adult mammals and by liver of fetal or neonatal mammals.

GENETICS  
#gene GDB:EPO  
#map\_position 7q21.3-q22.1  
#inons 5/1; 53/3; 82/3; 142/3  
CLASSIFICATION #superfamily erythropoietin  
KEYWORDS erythropoietin; glycoprotein; hormone; kidney; liver  
1-27  
28-193 #domain signal sequence #status predicted #label SIGV  
34-188, 56-60 #product erythropoietin #status experimental #label MARY  
51, 65, 110 #disulfide bonds #status experimental  
#binding\_site carbohydrate (Aan) (covalent) #status experimental  
153 #binding\_site carbohydrate (Ser) (covalent) #status experimental

SUMMARY #length 193 #molecular-weight 21307 #checksum 6371

Query Match 5.7%; Score 141; DB 2; length 193;  
Best Local Similarity 27.6%; Pred. No. 1.19e-03;  
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3;

Db 12 LILLELLEPLGLPGLCAPRLLCDSEVLYLLEAKAENITGCAEHLNITPTD 71  
||| : ||| : | : | : || ||| : : : : | : | : : | :  
Qy 7 LLVVMLLTARLT-LSSPAPACDLRLVLSKLIRDSVHLRSQCPVHP LPTVLLPA 65

Db 72 kmfyaawk-rmevvgqavevwqglalleaavllrgqall 108  
:: || : ||| : | : : ||| : : : ||| : : : ||| : : : |  
Qy 66 DFLSGEMKTQME-ETKADILLGAVTLLLEGVMAARGQL 102

RESULT 9  
ENTRY 9 #type complete  
TITLE LURB11  
ALTERNATE\_NAMES annexin XI - rabbit  
ORGANISM calyculin-associated annexin protein CAP-50  
#formal\_name Oryctolagus cuniculus #common\_name domestic rabbit  
DATE 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 24-Feb-1995

ACCESSIONS JH0694; PH0950; A38250; PS0263  
REFERENCE JH0694  
#authors Tokumitsu, H.; Mizutani, A.; Muramatsu, M.; Yokota, T.; Arai, K.; Hidaka, H.

	#journal	Biochem. Biophys. Res. Commun.	(1992)	186:	1227-1235
	#title	Molecular cloning of rabbit CAP-50, a calyculin-associated annexin protein.			
	#cross-references MUID:	J2378579			
	#accession	JH0694			
	#molecule_type	mRNA			
	#residues	1-503 ##label TOK			
	#cross-references DDBI:	D10883			
	#experimental_source	Lung			
	#accession	PH0950			
	#molecule_type	protein			
	#residues	104-141;213-231;254-262;270-280;285-309;319-337;429-448 478-492 ##label TR2			
REFERENCE	A38250	Tokumitsu, H.; Mizutani, A.; Minami, H.; Kobayashi, R.; Hidaka, H. J. Biol. Chem. (1992) 267:8919-8924 A calyculin-associated protein is a newly identified member of the Ca(2+)/phospholipid-binding proteins, annexin family.			
	#cross-references MUID:	J2250478			
	#accession	A38250			
	#molecule_type	protein			
	#residues	104-141;213-223,'X',225-231;254-263;271-280;285-291,'X', 293-300,'X',302-309;319-337;429-448;478-492 ##label T03			
COMMENT	This protein binds specifically to calyculin in a Ca2+ dependent manner.				
CLASSIFICATION	#superfamily	annexin VII; annexin repeat homology calcium binding; duplication; endonexin fold; glycoprotein; phospholipid binding			
KEYWORDS					
FEATURE					
	201-272	#domain annexin repeat homology #label AX1\			
	212-228	#region endonexin fold #status predicted\			
	273-344	#domain annexin repeat homology #label AX2\			
	284-300	#region endonexin fold #status predicted\			
	356-428	#domain annexin repeat homology #label AX3\			
	366-384	#region endonexin fold #status predicted\			
	432-503	#domain annexin repeat homology #label AX4\			
	443-459	#region endonexin fold #status predicted\			
	58	#binding site carbohydrate (Asn) (covalent) #status Predicted			
SUMMARY	#length	503 #molecular-weight	54034 #checksum	4870	
Query Match	5.6%; Score	139; DB 2;	Length	503;	
Best Local Similarity	25.7%;	Pred. No.	1.94e-03;		
Matches	26;	Conservative	30;	Mismatches	41;
				Indels	4;
				Gaps	4;
Db	36 glcdrvanygqfndqylsgmaamsqtfganvprnllypqapq-ggyvrpppygfqgppp	94			
Oy	233 SLDDIGCYLMRI-HELLINGRDLPPGFSRRITGCARDISSDTGSPLPMDRGYS-PSP	296			
Db	95 tqgsrvppvygvrrpppggnppsgyspsypptfgavrypqgmppp	135			
Oy	297 THPRGTQVTLFRLPTLFRVVQLHLRPDPAFL-PTPRS	336			
RESULT	10				
ENTRY TITLE	A33647				
ORGANISM	sulfated surface glycoprotein 185 - Volvox carteri				
DATE	11-Apr-1990 #sequence_revision	11-Apr-1990 #text_change			
ACCESSIONS	A33647				

[illegible]

```
Qy 7 LUVWMLLTARLTLSPPAPP-ACDRLVLSKLRDSDVLSRSLSCPEVHPLPVPVLLPAAV 65
Db 72 kmfyawk-rmevgqgavewgqllaleavlrqgavla 109
:: || || | :: ::|| |:: ||:
#cross-references ETKAQDILGAVTLLEGVMAARGQLG 103

RESULT 12
ENTRY A24901 #type complete
TITLE erythropoietin precursor - mouse
ORGANISM #formal name Mus musculus #common name house mouse
DATE 19-May-1989 #sequence_revision 19-May-1989 #text_change
18-Jun-1993

ACCESSIONS A24901
REFERENCE A24901
#authors McDonald, J.D.; Lin, F.K.; Goldwasser, E.
#journal Mol. Cell. Biol. (1986) 6:842-848
#title Cloning, sequencing, and evolutionary analysis of the mouse
#cross-references MUID:87039104
#accession A24901
#molecule_type mRNA
#residues 1-193 ##label MCD
CLASSIFICATION #superfamily erythropoietin
KEYWORDS erythropoiesis; glycoprotein; hormone
SUMMARY #length 193 #molecular-weight 21223 #checksum 6523

Query Match 5.1%; Score 127; DB 5; Length 193;
Best Local Similarity 25.3%; Pred. No. 3.41e-02;
Matches 25; Conservative 30; Mismatches 41; Indels 3; Gaps 3;

Db 12 llslvelplgipygpprlcdervlyllekeavntmgseesclneltvptd 71
||: || | | ::|| | || | | :: | | : | |
Qy 7 LUVWMLLTARLTLSPPAPP-ACDRLVLSKLRDSDVLSRSLSCPEVHPLPVPVLLPAAV 65
Db 72 kmfyawk-rmevgqgavewgqllaleavlrqgavla 109
:: || || | :: || | :: || | :: || | :: || |
Qy 66 DFLGSEKKTQME-ETKAQDILGAVTLLEGVMAARGQLG 103

RESULT 13
ENTRY LUBO11 #type complete
TITLE annexin XI form A - bovine
ALTERNATE_NAMES calyculin-associated protein peptide, CAP-50
ORGANISM #formal name Bos primigenius taurus #common name cattle
DATE 31-Mar-1993 #sequence_revision 31-Mar-1993 #text_change
23-Mar-1995

ACCESSIONS A42113; A42909; B42909; C42909; D42909
REFERENCE A42113
#authors Towle, C.A.; Treadwell, B.V.
#journal J. Biol. Chem. (1992) 267:5416-5423
#title Identification of a novel mammalian annexin. cDNA cloning,
sequence analysis, and ubiquitous expression of the annexin
XI gene.
#cross-references MUID:92184796
#accession A42113
#molecule_type mRNA
#residues 1-503 ##label TOW
#cross-references GB:M82802
#note the authors did not translate the codon GAC for residue
503

REFERENCE A42909
#authors Mitutani, A.; Ueda, N.; Tokumitsu, H.; Minami, H.; Yasui,
K.; Kobayashi, R.; Hidaka, H.
```

```
#journal J. Biol. Chem. (1992) 267:13498-13504
#title CAP-50, a newly identified annexin, localizes in nuclei of
cultured fibroblast 3Y1 cells.
#cross-references MUID:92317074
#accession A42909
#molecule_type protein
#residues 213-223,'X', 225-226;319-325,'G', 327-328,'X', 330-333,'Q',
335-339;441-450,'X', 452-454,'D', 456-460;479-497
##label MITZ
#experimental_source lung
#note Annexins undergo reversible, calcium-dependent binding to membrane
phospholipids. Although a number have been proposed, the
physiological roles of the various annexins are not yet fully
understood.

GENETICS 19/1; 56/2
#introns the list of introns is incomplete
#note the list of introns is incomplete
CLASSIFICATION #superfamily annexin VII; annexin repeat homology
KEYWORDS calcium binding; duplication; endonexin fold; glycoprotein;
phospholipid binding

FEATURE 201-272
212-228 #region endonexin fold #status predicted\
273-344 #domain annexin repeat homology #label AX2\
284-300 #region endonexin fold #status predicted\
356-428 #domain annexin repeat homology #label AX3\
368-384 #region endonexin fold #status predicted\
432-503 #domain annexin repeat homology #label AX4\
443-459 #region endonexin fold #status predicted\
59,111 #binding site carbohydrate (Aan) (covalent) #status
predicted

SUMMARY #length 503 #molecular-weight 54018 #checksum 5855

Query Match 5.1%; Score 126; DB 2; Length 503;
Best Local Similarity 23.8%; Pred. No. 4.30e-02;
Matches 24; Conservative 29; Mismatches 44; Indels 4; Gaps 4;

Db 37 qldhvanvyaqgfngdylegvaamsqtfqga-nvnpillypqa9-gyvpvpgqf9qpp 94
||: || | | ::|| | | | : | | : ||: || | |
Qy 233 SLDDIPETLARI-HELLNGTGLGFPGRSRRTLGAPDISSGTSIDTGSILPRLDPTGS-PSP 296
Db 95 aqgpyvpygmypppgpnprsgmpeyppvpgapvpgqmlpp 135
:: | | | | : | | | | | | | | | | |
Qy 297 TTPPTGQYTLFPLPPLPPVQVHLPLDPSAPTPPTPS 337

RESULT 14
ENTRY S22456 #type complete
TITLE hydroxyproline-rich glycoprotein - perenial teosinte
ORGANISM #formal name Zea diploperennis #common name perenial
teosinte
DATE 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change
18-Jun-1993

ACCESSIONS S22456
REFERENCE S22456; S18964
#authors Raz, R.; Jose, M.; Moya, A.; Martinez-Iguierdo, J.A.;
Pulidomenech, P.
#journal Mol. Gen. Genet. (1992) 233:252-259
#title Different mechanisms generating sequence variability are
revealed in distinct regions of the hydroxyproline-rich
glycoprotein gene from maize and related species.
#accession S22456
#molecule_type DNA
```



(TM)

Release 2.1D John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993, 1994, 1995 University of Edinburgh, U.K.  
Distribution rights by IntelliGenetics, Inc.

MParch\_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Apr 1 11:00:24 1996; MasPar time 7.11 Seconds

Tabular output not generated.

Title: >US-08-347-748-2

Description: (1=353) from US08347748.pep

Sequence: 1 MELTELLVMLLTARLT.....PTSP LINTSYTHSQNLSEQ 353

Scoring table: PAM 150

## Gap 11

Searched: 43470 seqs, 15335248 residues

Post-processing: Minimum Match 0%

### Listing first 45 summaries

Database: swiss-prot31

1:part1 2:part2 3:part3 4:part4 5:part5 6:part6 7:part7  
8:part8

Statistics: Mean 48.744; Variance 109.173; scale 0.446

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result	No.	Score	Query Match	Length	DB	ID	Description	Pred. No
1	2493	100.0	35.3	7	TPO_HUMAN		THROMBOPOIETIN PRECUR	0.00e+00
2	1748	70.1	35.6	7	TPO_MOUSE		THROMBOPOIETIN PRECUR	2.33e-29
3	162	6.5	192	3	EPO_MOUSE		ERYTHROPOIETIN PRECUR	3.81e-08
4	156	6.3	188	3	EPO_FELCA		ERYTHROPOIETIN PRECUR	3.58e-07
5	155	6.2	175	3	EPO_CANFA		ERYTHROPOIETIN PRECUR	3.56e-07
6	154	6.2	192	3	EPO_RAT		ERYTHROPOIETIN PRECUR	4.84e-07
7	141	5.7	193	3	EPO_HUMAN		ERYTHROPOIETIN PRECUR	2.67e-05
8	139	5.6	503	1	ANXA_RABIT		ANNEKIN X1 (CALYCICLIN	4.88e-05
9	138	5.5	194	3	EPO_SHEEP		ERYTHROPOIETIN PRECUR	6.59e-05
10	132	5.3	485	7	S5SG_YOLCA		SULFATED SURFACE GLYC	3.99e-04
11	127	5.1	192	3	EPO_MACFA		ERYTHROPOIETIN PRECUR	1.66e-03
12	126	5.1	503	1	ANXA_BOVIN		ANNEKIN X1 VARIANT 1	2.22e-03
13	116	4.7	751	8	VAP1_YEAST		VERPROLIN.	3.63e-02

15	14	115	4.6	441	8	YK92	YEAST	HYPOTHETICAL 49.6 KD	4.76e-02
14	115	4.6	1874	6	POLR_KRYM2	RNA REPLICASE POLYPRO	4.76e-02	1.07e-01	
13	115	4.6	267	3	EXN1_MAIZE	EXTENSIN PRECURSOR (P	1.07e-01	1.07e-01	
16	112	4.5	1232	8	Y003_CAEEL	HYPOTHETICAL 134.9 KD	1.07e-01	1.07e-01	
17	112	4.5	1232	8	Y003_CAEEL	HYPOTHETICAL 134.9 KD	1.07e-01	1.07e-01	
18	110	4.4	153	8	Y016_ECOLI	HYPOTHETICAL 16.2 KD	2.36e-01	2.36e-01	
19	109	4.4	474	8	VT93_TTIV1	VIRAL PROTEIN TPX...	2.36e-01	2.36e-01	
20	109	4.4	505	1	ANXB_BOVIN	ANNEKIN XI VARIANT 2	2.36e-01	2.36e-01	
21	109	4.4	626	3	GRB1_HUMAN	PLATELET G1XOPROTEIN	2.36e-01	2.36e-01	
22	106	4.3	153	8	YH17_YEAST	HYPOTHETICAL 17.1 KD	5.14e-01	5.14e-01	
23	106	4.3	323	3	HN33_MOUSE	HEPAROXYTE NUCLEAR FA	5.14e-01	5.14e-01	
24	106	4.3	1206	3	FOR4_MOUSE	FORMIN 4 (LIMB DEFORM	5.14e-01	5.14e-01	
25	106	4.3	1468	3	FORN_MOUSE	FORMIN (LIMB DEFORMAT	5.14e-01	5.14e-01	
26	105	4.2	620	3	EXT1_TOBAC	EXTENSIN PRECURSOR (C	6.64e-01	6.64e-01	
27	105	4.2	742	1	BAL1_HUMAN	BILE-SALT-ACTIVATED 1	6.64e-01	6.64e-01	
28	105	4.2	759	3	GSPD_XANOP	GENERAL SECRETION PAT	6.64e-01	6.64e-01	
29	103	4.1	354	4	HN33_RAT	HEPAROXYTE NUCLEAR FA	1.10e+00	1.10e+00	
30	103	4.1	354	4	LAC1_KLEPN	LACTOSE OPEIRON REBERS	1.10e+00	1.10e+00	
31	102	4.1	421	1	ACRO1_HUMAN	ACROSIN PRECURSOR (EC	1.42e+00	1.42e+00	
32	102	4.1	437	1	ACRO1_RAT	ACROSIN PRECURSOR (EC	1.42e+00	1.42e+00	
33	102	4.1	449	6	RRX8_MOUSE	RETINOIC ACID RECEPT	1.42e+00	1.42e+00	
34	101	4.1	1844	6	POLR_TYMW	RNA REPLICASE POLYPRO	1.82e+00	1.82e+00	
35	101	4.1	1844	6	POLR_TYMW	RNA REPLICASE POLYPRO	1.82e+00	1.82e+00	
36	102	4.1	3079	4	IRAZ_YEAST	INHIBITORY REGULATOR	1.42e+00	1.42e+00	
37	100	4.0	312	7	US10_HSV1	VIRION PROTEIN US10.	2.33e+00	2.33e+00	
38	99	4.0	312	3	GDF1_HUMAN	EMBRIONARY GROWTH FACT	2.98e+00	2.98e+00	
39	99	4.0	496	4	HXA4_HUMAN	HOMEODOM PROTEIN HOX-	2.98e+00	2.98e+00	
40	100	4.0	534	1	APG_ABAFH	ANTER-SPECIFIC PROLIN	2.33e+00	2.33e+00	
41	99	4.0	543	3	F1BF_ADECG	FIBER PROTEIN.	2.98e+00	2.98e+00	
42	100	4.0	1447	7	TPS1_YEAST	HELICASE TP51.	2.33e+00	2.33e+00	
43	100	4.0	2274	5	MUC2_HUMAN	MUCIN 2 (INTESTINAL, M	2.33e+00	2.33e+00	
44	98	3.9	187	8	YK93_YEAST	HYPOTHETICAL 20.2 KD	3.80e+00	3.80e+00	
45	98	3.9	657	8	YK93_YEAST	HYPOTHETICAL 12.7 KD	3.80e+00	3.80e+00	

## ALIGNMENTS

ID	RESULT	1	STANDARD;	PRT;	353 AA.
AC	P40725;	DT	01-FEB-1995 (REL. 31, CREATED)		
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)	DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)		
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)	DE	THROMBOPELTIN PRECURSOR (MEGAKARYOCYTE COLONY STIMULATING FACTOR)		
DE	(C-MPL LIGAND) (ML) (MEGAKARYOCYTE GROWTH AND DEVELOPMENT FACTOR)	DE	(MGDF) .		
GN	THPO.	OS	HOMO SAPIENS (HUMAN) .		
OC	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	OC	EUTHERIA; PRIMATES.		
OC	EUTHERIA; PRIMATES.	RN	(1)		
RN	(1)	RP	SEQUENCE FROM N.A.		
RP	94261202	RL	(2)		
RA	DE SAUVAGE F.J., HASS P.E., SPENCER S.D., MALLOY B.E., GUREY A.L.,	RA	SPENCER S.A., DARBOONE W.C., HENZEL M.J., WONG S.C., KUANG M.-J.,		
RA	OLDS K.J., HULTGREN B., SOLBERG L.A. JR., GOEDEL D.V., EATON D.L.,	RA	NATURE 369:533-538(1994) .		
RL	(2)	RN	SEQUENCE FROM N.A.		
RP	94291201	RC	TISSUE=LIVER;		
RA	BARTLEY T.D., BOGENBERGER J., HUNT P., LI Y.S., LU H.S.,	RA	MARTIN F., CHANG M.S., SAMAL B.B., NICHOL S., BOSSELMAN R.A.,		
RL	CELL 77:1117-1124(1994) .	CC	-I- FUNCTION: ACTS AS A CIRCULATING REGULATOR OF PLATELET NUMBERS.		

CC	MEGACARBOCYTIC LINEAGE SPECIFIC GROWTH AND DIFFERENTIATION FACTOR.
CC	-1- SUBCELLULAR LOCATION: SECRETED.
CC	EMBL; L33410; HSMICPL.
DR	EMBL; U11025; HS11025.
DR	PIR; S45331; S45331.
DR	MIW; 600044; 11TH EDITION.
KM	GLYCOPROTEIN; HORMONE; SIGNAL.
FT	SIGNAL 1 21
FT	CHAIN 22 353
FT	CARBOHYD 197 197
FT	CARBOHYD 206 206
FT	CARBOHYD 234 234
FT	CARBOHYD 255 255
FT	CARBOHYD 340 340
FT	CARBOHYD 348 348
FT	CONFLICT 113 113
SO	SEQUENCE 353 AA; 37822 MW; 702923 CDF; Q -> E (IN REF. 2).

Query Match	100.0%;	Score 2493;	DB 7;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 0.00e+00;		
Matches 353; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Db	1	metellllymlllratcltclspapaccdryrlakllrdshvherlrlsqcwpnlpjrv	60
Qy	1	MELTELLLYMVLMLLTATLTLTSSPAPACDRLVSLTLRLDSHVHLHSLSQCEVHPRLPV	60
Db	61	llpavfslgawkqmeetkagdlvgavrlllgymaargqljptclslllgqlsgqvrll	120
Qy	61	LIPAVDFSIGEMKQMEETKAGDILGAVTLLLEGVMAARGQLPTCLSLILGQSQVRL	120
Db	121	llgalsslglgtcjprrgqrrtahnkpnafisfqhlllgkvrflmlvggselcvrrappt	180
Qy	121	LLGALGSLGSLGTQJPRPGQRTAHKPNMFIISQGHLLRKKVRFMLVGSGSLTCVRRAPPT	180
Db	241	avpertsivtlcmlnlpnrtsqllletinfcaartctsglllkwqgqfrakipjllntrsrsl	240
Qy	181	AVPSRISLVLTLMELNPKRISGLLEINFATASATTSGLLKWQGGFRAKITGELLNQSLSL	240
Db	301	tdqipylnrtheellntrgrllfpqpserrcllgandiesqtsdgselpnllqgvyaparrthp	300
Qy	241	DQIPGILNRHIELLNTRGLFPGPSRRTTGCAADISSGTSQTSLEPNLDPQVSPSPTHPP	300
Db	301	tggvclfpjlpjrcldptpvgdlhlpdpasapctpcpgrllntsythsmlesqeg	353
Qy	301	TGGVTLFPJLPJRLPTPVQDLHLDPDASAPCTPCPGRLLNTSYTHSMLESQEG	353

RA	RESULT	2
AC	ID	TPO MOUSE STANDARD; PRT; 356 AA.
AC	P40226;	
DT	01-FEB-1995 (REL. 31, CREATED)	
DT	01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)	
DT	01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)	
DE	THROMBOPOIETIN PRECURSOR (MEGAKARYOCYTE COLONY STIMULATING FACTOR)	
DE	(C-MPL LIGAND) (ML) (MEGAKARYOCYTE GROWTH AND DEVELOPMENT FACTOR)	
DE	(MODE).	
DE	MUS MUSCULUS (MOUSE).	
OC	EDUARDO A. REINZOLA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;	
OC	EUTHERIA; RODENTIA.	
OC	[1]	
RP	SEQUENCE FROM N.A.	
RM	94261207	
RA	LOK S., KUSHANSKY K., HOLLY R.D., KUIPERS J.L., LOFTON-DAY C.E.,	

RA OORT P.V., GRANT F.J., HEIHEL M.D., BURKHEAD S.K., KRAMER J.M.,  
RA BELL I.A.N., SPECHER C.A., BLUMBERG H., JOHNSON R., PRINARD D.,  
RA CHING A.F.T., MATTHEWS S.L., BAILEY M.C., FOUSTON J.W., BIDDLE M.M.,  
RA OSBORNE S.G., EVANS S.J., SHEPPARD P.O., PRESNELL S.R., O'HARA P.J.,  
RA HAGEN F.S., ROTH G.J., FOSTER D.C.;  
RL NATURE 369:565-568 (1994).  
CC -I- FUNCTION: ACTS AS A CIRCULATING REGULATOR OF PLATELET NUMBERS.  
CC MEGAKARYOCYTIC LINEAGE SPECIFIC GROWTH AND DIFFERENTIATION  
CC FACTOR.

CC	-1-	SUBCELLULAR LOCATION: SECRETED.
DR	EMBL, L34169; MATPROA.	
DR	PIR; S45330; S45330.	
KM	GLYCOPROTEIN; HORMONE; SIGNAL.	
FT	SIGNAL	1 21 POTENTIAL.
FT	CHAIN	22 356 THROMBOPOIETIN.
FT	CARBOHYD	197 197 POTENTIAL.
FT	CARBOHYD	206 206 POTENTIAL.
FT	CARBOHYD	235 235 POTENTIAL.
FT	CARBOHYD	249 249 POTENTIAL.
FT	CARBOHYD	256 256 POTENTIAL.
FT	CARBOHYD	336 336 POTENTIAL.
FT	CARBOHYD	351 351 POTENTIAL.
Q0	SEQUENCE	356 AA; 701294 CN;

Query Match	70.1%;	Score 1748;	DB 7;	Length 356;
Best Local Similarity	71.6%;	Pred. No. 2.32e-294;		
Matches 255; Conservative	47;	Mismatches 49;	Indels 5;	Gaps 4;

[illegible]

RESULT	3			
OS	ID	EPO MOUSE	STANDARD;	PRT, 192 AA.
CC	AC	P07321;		
OC	DT	01-APR-1968 (REL. 07, CREATED)		
OC	DT	01-APR-1968 (REL. 07, LAST SEQUENCE UPDATE)		
DT	DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)		
DE		ERYTHROPOIETIN PRECURSOR.		
GN	EPO.			
OS	MUS MUSCULUS (MOUSE).			
CC	EDUARDOYA, METAZOA;			
OC	CHORDATA, VERTEBRATA, TETRAPODA,			
OC	MAMMALIA,			
OC	EUTHERIA, RODENTIA.			



[illegible]

	CC	-I-	FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
	CC	-I-	TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
	CC	-I-	SUBCELLULAR LOCATION: SECRETED.
	DR	EMBL; LI0606;	FCEERYTHO.
	DR	PROSITE; PS00817;	EPO.
	KW	ERYTHROCYTE MATURATION;	GLYCOPROTEIN; HORMONE; SIGNAL.
	FT	NON TER	1 1
	FT	SIGNAL	<1 22 BY SIMILARITY.
	FT	CHAIN	23 188 ERYTHROPOIETIN.
	FT	DISELFD	29 183 BY SIMILARITY.
	FT	DISELFD	51 55 BY SIMILARITY.
	FT	CARBONYD	46 46 POTENTIAL.
	FT	CARBONYD	60 60 POTENTIAL.
	FT	CARBONYD	105 105 POTENTIAL.
	SQ	SEQUENCE	188 AA; 20608 MM; 175419 CN;
	Query Match	6.3%; Score 156;	DB 3; Length 188; Best Local Similarity 24.8% Pred. No. 2.58e-07;
	Matches	36; Conservative	45; Mismatches 59; Indels 5; Gaps 5;
Db	6	l  lllpjqlpvlgqaprlldicarfvlrylleareaeenymvgcagcsfeenttynd	65
Oy	6	:      l :   ::         : : : : : : : : : :   : l : : : : :	64
LIIIVMLLTATRLT-LISDPAPACDLVLSKLIRSHYLSRSLQCPEVHPLPVLPA			
Db	66	tkmfvfytv-imdyggqvavewqglalalseallrqgalanssqpaetqlghvkavasl	124
Oy	65	VDFSLSEKTKQE-EETADDIIGAVTTLLLEGVMARGQLPTCLSSLGQSGLVRIL-L	122
:	: ::  :	:>::   :: :	:
Db	125	rslselilra-lgaqkaeatlepeats	148
Oy	123	GALOSLIQTQLPPGGRITTHKPDA	147
RESULT	5	STANDARD;	PRT; 175 AA.
ID	EPO CANFA		
AC	P33707;		
DT	01-FEB-1994 (REL. 28, CREATED)		
DT	01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)		
DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)		
DE	ERYTHROPOIETIN PRECURSOR (FRAGMENT).		
GN	EPO.		
OC	CANIS FAMILIARIS (DOG).		
OS	EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;		
OC	EUTHERIA; CARNIVOA.		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RM	93372347		
RA	WHEN D., BOISSEL J.P.R., TRACY T.E., MULCAHY L.S., CZELUSNIAN J.,		
RA	GOODMAN M., BURN H.F.;		
RU	BLOOD 82:1507-1516(1993).		
CC	-I-	FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.	
CC	-I-	TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.	
CC	-I-	SUBCELLULAR LOCATION: SECRETED.	
CC	EMB1; LI3027;	FCEERYPB.	
DR	PROSITE; PS00817;	EPO.	
FM	ERYTHROCYTE MATURATION;	GLYCOPROTEIN; HORMONE; SIGNAL.	
KT	NON_TER	1 1	
FT	SIGNAL	<1 22 BY SIMILARITY.	



STRUCTURE OF CARBOHYDRATES.

RA TAKEDUCHI M., KOBATA A.;  
RM 92314463  
GLYCOBIOLOGY 1:337-346(1991).

-|- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.

-|- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS AND BY LIVER OF FETAL OR NEONATAL MAMMALS.

-|- SUBCELLULAR LOCATION: SECRETED.  
CC EMBL; X02158; HSERRG.  
DR EMBL; X02157; HSEPRP.  
DR EMBL; M1319; HSERPA.  
DR PIR; A01855; ZUHU.  
DR PIR; A25384; A25384.  
DR PIR; A24744; A24744.  
DR PIR; A22210; A22210.  
DR MIM; 133170; 11TH EDITION.  
DR PROSITE; PS00817; EPO.  
KM ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.

FT SIGNAL 1 27  
FT CHAIN 28 193  
FT DISULEID 34 188 ERYTHROPOIETIN.  
FT DISULEID 56 60  
FT CARBOHYD 51 51  
FT CARBOHYD 65 65  
FT CARBOHYD 110 110  
FT CARBOHYD 153 153  
FT PROPEP 190 193 MAY BE REMOVED IN PROCESSED PROTEIN.  
FT CONFLICT 85 85 Q -> QQ (IN REF. 3).  
SQ SEQUENCE 193 AA; 21306 MW; 182311 CN;

Query Match 5.7%; Score 141; DB 3; Length 193;  
Best Local Similarity 27.6%; Pred. No. 2,67e-05;  
Matches 27; Conservative 27; Mismatches 41; Indels 3; Gaps 3.

Db 12 llllellelplgplvlgpprlcdervyleylleakeanittgcacnchsenitcypdt 71  
|| : | | | | : || : || : | : : : | : | : | : | : |  
Qy 7 LVWVLLTARLT-ISSPAPACDLRLVLSKLRDSHVLHSRLSGCEVHPLPPTVLLPAV 65

Db 72 kmfyawk-rmevgqagvevwgqlallseavlrpqall 108  
:: | | : | : : : || : | : | : | : | : | : | : |  
Qy 66 DFLSLEMKTKOME-ETKAQDILCAVTVLLIECVMAANGQL 102

RESULT 8  
ID ANYA RABIT STANDARD; PRT; 503 AA.  
AC P33477;  
DT 01-FEB-1994 (REL. 28, CREATED)  
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)  
DE ANNEXIN XI (CALCICLIN-ASSOCIATED ANNEXIN 50) (CAP-50).  
OS ORCOTOLAGUS CINCTICULUS (RABBIT).  
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
CC EUTHERIA; LACOMORPHA.  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TISSUE=LUNG;  
RM 92378579  
RA TOKUMITSU H., MIZUTANI A., MORIMATSU M., YOKOTA T., ARAI K.,  
RA HIDAKA H.;  
RL BIOCHEM. BIOPHYS. RES. COMMUN. 186:1227-1235(1992).  
CC -|- FUNCTION: BINDS SPECIFICALLY TO CALCICLIN IN A CALCIUM-DEPENDENT MANNER.

```
CC      -|- DOMAIN: CONTAINS FOUR HOMOLOGOUS REPEATS WITH A CONSENSUS  
CC      SEQUENCE COMMON TO ALL ANNEKIN PROTEINS. A PAIR OF THESE REPEATS  
CC      MAY FORM ONE BINDING SITE FOR CALCIUM AND PHOSPHOLIPID.  
CC      -|- SIMILARITY: TO OTHER PROTEINS OF THE ANNEKIN FAMILY.  
DR      EMBL; D10883; OCCAP.  
DR      PIR; JH0694; UBRB1.  
DR      HSSP; P19999; ICLG.  
DR      PROSITE; PS00223; ANNEKIN.  
KM      ANNEKIN; CALCIUM/PHOSPHOLIPID-BINDING; REPEAT.  
FT      REPEAT       207   267     ANNEKIN.  
FT      REPEAT       279   339     ANNEKIN.  
FT      REPEAT       363   423     ANNEKIN.  
FT      REPEAT       438   498     ANNEKIN.  
SQ      SEQUENCE    503 AA; 54034 MW; 1231578 CN;  
  
Query Match          5.6%; Score 139; DB 1; Length 503;  
Best Local Similarity 25.7%; Pred. No. 4,8e-05;  
Matches    26; Conservative    30; Mismatches 41; Indels    4; Gaps    4;  
  
Db      36 glhdvanyagfndyldgmaamsgtfgganvpnnlypaapq-ggyppvpggfgqpp 94  
QY      239 SLDDIPEVYNLR-HELLNGTRGLRFGPSRRRLTCAVDISSGTSIGSLPNLDPEYS-PSP 296  
DB      95 tqdvepppygvypgpnmgpvgysppfgapypvqqmppp 135  
QY      297 THPTGTGYTLFPLPTLPFPVVQLHLPLDPSSAPT-PPPTS 336  
  
RESULT      9  
ID      EPO SHEEP      STANDARD;      PRT;      194 AA.  
AC      P33709;  
DT      01-FEB-1994 (REL. 28, CREATED)  
DT      01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)  
DT      01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)  
DE      ERYTHROPOIETIN PRECURSOR.  
GN      EPO.  
OS      OVIS ARIES (SHEEP).  
OC      EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;  
OC      EUTHERIA; ARTIODACTYLIA.  
RN      [1]  
RP      SEQUENCE FROM N.A.  
RC      TISSUE=KIDNEY;  
RM      93351736  
FU      FU P., EVANS B., LIM G.B., MORITZ K., WINTOUR M.E.;  
RL      MOL. CELL. ENDOCRINOLOGY. 93:107-116(1993).  
CC      -|- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE  
CC      REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A  
CC      PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.  
CC      -|- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS  
CC      AND BY LIVER OF FETAL OR NEONATAL MAMMALS.  
CC      -|- SUBCELLULAR LOCATION: SECRETED.  
DR      EMBL; Z24681; OAEERYOIA.  
DR      PROSITE; PS00817; EPO.  
KM      ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.  
FT      SIGNAL        1      27     BY SIMILARITY.  
FT      CHAIN         28      194   ERYTHROPOIETIN.  
FT      DISULFID      34      189   BY SIMILARITY.  
FT      DISULFID      56      60     BY SIMILARITY.  
FT      CARBOHYD      51      51     POTENTIAL.  
FT      CARBOHYD      65      65     POTENTIAL.  
FT      CARBOHYD      110     110    POTENTIAL.  
SQ      SEQUENCE    194 AA; 21335 MW; 175348 CN;
```

Db	11	f l p g l p v a p p l l c a r l e a a e a n a t g c a g c f a e n t y p d	70
Db	11	f l p g l p v a p p l l c a r l e a a e a n a t g c a g c f a e n t y p d	70
Qy	6	LLLVALLLTTLRLT-LSSPAPPCADLNLSTKLRSVHLSRUSQCEVHPLEPPVLLPA	64
Db	71	tknfyak-rmevqgalewqgallaeatrgalla	109
Qy	65	VDESLGEMKTOME-ETKADIIAGVATLLLEGVAAKQGLG	103
RESULT	10		
ID	SSGP VOLCA	STANDARD;	PRT; 485 AA.
AC	P2197;		
DT	01-AUG-1991 (REL. 19, CREATED)		
DT	01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)		
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)		
DE	SULFATED SURFACE GLYCOPROTEIN 185 (SSG 185).		
OS	VOLVOX CARTERI.		
OC	EUKARYOTA; PLANTA; PHYCOPHYTA; CHLOROPHYTA (GREEN ALGAE);		
OC	CHLOROPHYCEAE; VOLVOCALES; VOLVOCAEAE.		
RA	ERTL H., MENGELE R., WENZL S., ENGEL J., SUMPER M.;		
RM	J. CELL. BIOL. 109:3493-3501(1989).		
CC	-I- FUNCTION: THE EXTRACELLULAR MATRIX (ECM) OF VOLVOX CONTAINS		
CC	INSOLUBLE FIBROUS LAYERS THAT SURROUND INDIVIDUAL CELLS AT A		
CC	DISTANCE TO FORM CONTIGUOUS CELLULAR COMPARTMENTS. SSG 185 IS THE		
CC	MONOMERIC PRECURSOR OF THIS SUBSTRUCTURE (C32 STRUCTURE). THE		
CC	COVALENT CROSS-LINKS ARE FORMED BETWEEN THE SACCHARIDE CHAINS		
CC	RATHER THAN BETWEEN THE POLYPEPTIDE CHAINS.		
CC	-I- PTM: A CENTRAL DOMAIN OF 77 AA CONSISTS ALMOST EXCLUSIVELY OF		
CC	HYDROXYPROLINE RESIDUES.		
DR	PIR; A33647; A33647.		
DR	HSSP; P19999; ICLG.		
KM	GLYCOPROTEIN; SULFATATION; HYDROXYLATION.		
FT	DOMAIN 228 340 PRO-RICH.		
FT	DOMAIN 260 293 POLY-PRO.		
FT	SEQUENCE 485 AA; 50436 MW; 1322558 CN;		
Qy	269	LGAPDLS-G-TSITGSLPPIQLQVSPSTHPPGQITLLEPPLTLPVQVQLHPLPD	326
Db	271	pppppppppp 281	
Qy	327	PSAPPTPTSP 337	
RESULT	11		
ID	EPO MACFA	STANDARD;	PRT; 192 AA.
AC	P07865;		
DT	01-AUG-1988 (REL. 08, CREATED)		
DT	01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)		
DT	01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)		
DE	ERYTHROPOIETIN PRECURSOR.		
EN	EPO.		

```

0S MACACA FASCICULARIS (GRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; PRIMATES.
RN (1)
RP SEQUENCE FROM N.A.
RM 87055236
RA LIN F.-K., LIN C.-H., LAI P.-H., BROWNE J.K., EGRIE J.C., SMALLING R.,
RA FOX G.M., CHEN K.K., CASTRO M., SUGGS S.;
RL GENE 44:201-209(1986).
CC -|- FUNCTION: ERYTHROPOIETIN IS THE PRINCIPAL HORMONE INVOLVED IN THE
CC REGULATION OF ERYTHROCYTE DIFFERENTIATION AND THE MAINTENANCE OF A
CC PHYSIOLOGICAL LEVEL OF CIRCULATING ERYTHROCYTE MASS.
CC -|- TISSUE SPECIFICITY: PRODUCED BY KIDNEY OR LIVER OF ADULT MAMMALS
CC AND BY LIVER OF FETAL OR NEONATAL MAMMALS.
CC -|- SUBCELLULAR LOCATION: SECRETED.
DR EMBL; M18189; MFEP.
DR PIR; JQ0173; JQ0173.
KM PROSITE; PS00817; EPO.
KR ERYTHROCYTE MATURATION; GLYCOPROTEIN; HORMONE; SIGNAL.
FT SIGNAL 1 27 BY SIMILARITY.
FT CHAIN 28 192 ERYTHROPOIETIN.
FT DISULFID 34 187 BY SIMILARITY.
FT DISULFID 56 60 BY SIMILARITY.
FT CARBOHYD 51 51 BY SIMILARITY.
FT CARBOHYD 65 65 BY SIMILARITY.
FT CARBOHYD 110 110 BY SIMILARITY.
FT CARBOHYD 152 152 BY SIMILARITY.
SQ SEQUENCE 192 AA; 21113 MW; 175216 CN;

Query Match 5.1%; Score 127; DB 3; Length 192;
Best Local Similarity 25.3%; Pred. No. 1,66e-03;
Matches 25; Conservative 30; Mismatches 41; Indels 3; Gaps 3

Db 12 |||l||v||p||g||p||g||p||c||d||c||r||v||l||y||l||e||k||e||a||n||t||v||n||g||c||s||c||s||e||n||t||y||p||d||t 71
   ||: :| | :|:::| | ||| :| :| :| :| :|
Qy 7 LIVVMLLTARLITSSAP-ACDIKRVSLKLSNDSHVLRHSLOCFEYHPLEPTVLPAV 65
Db 72 kmfyawk-rimevgqgvewvglla||eeav||rgava 109
   :: ||:|| |:::|||| |:| :| :| :| :|
Qy 66 DFLSGEKTKTOME-ETKADILGAVTLLEGVMARQGLG 103

RESULT 12
ID ANXA BOVIN STANDARD; PTR; 503 AA.
AC P27214;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE ANNEXIN XI VARIANT 1 (CALCICYCLIN-ASSOCIATED ANNEXIN 50) (CAP-50).
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC EUTHERIA; ARTIODACTYLIA.
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=CHONDROCYTE;
RM 92184796
RA TOWLE C.A., TREADWELL B.V.;
RL J. BIOL. CHEM. 267:5416-5423(1992).
RP (2)
RP SEQUENCE FROM N.A.
RC TISSUE=CARTILAGE;
RM 92305067
RA TOWLE C.A., WEISSBACH L., TREADWELL B.V.;
```

RL	BIOCHEM BIOPHYS. ACTA 1131:223-226(1992).				
CC	-1- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.				
CC	-1- DOMAIN: CONTRAINS FOUR HOMOLOGOUS REPEATS WITH A CONSENSUS				
CC	SEQUENCE COMMON TO ALL ANNEXIN PROTEINS. A PAIR OF THESE REPEATS				
CC	MAY FORM ONE BINDING SITE FOR CALCIUM AND PHOSPHOLIPID.				
CC	-1- SIMILARITY: TO OTHER PROTEINS OF THE ANNEXIN FAMILY.				
CC	-1- ALTERNATIVE PRODUCTS: TWO FORMS OF ANNEXIN XI ARE PRODUCED BY				
CC	ALTERNATIVE SPLICING OF THE SAME GENE.				
DR	EMBL, M82802; BTRNXXIS.				
DR	PIR, A42113; LUBO11.				
DR	HSSP; P19999; ICLG.				
DR	PROSITE, PS00223; ANNEXIN.				
KW	ANNEXIN; CALCIUM/PHOSPHOLIPID-BINDING; REPEAT; ALTERNATIVE SPLICING.				
FT	REPEAT 207 267 ANNEXIN.				
FT	REPEAT 279 339 ANNEXIN.				
FT	REPEAT 363 423 ANNEXIN.				
FT	REPEAT 438 498 ANNEXIN.				
SQ	SEQUENCE 503 AA; 54018 MW; 1237712 CN;				
Db	31 gldhvanvsgqfngdyLsgvaamgtfgyga-nvnlYpqaqg-9gypvppgfgqppp 94				
Qy	238 SLIDITGYLNRI-HELLINGRCLFPGFSRRTLCADPDISGSDTGSPLNIDPetyS-psp 296				
Db	95 aqdvpsygmypppggnpcsgmppsppypqapvpgqmlpp 135				
Qy	297 THTPTGQYTLFPLPPLTPVVLHPILPPLPSAPPTPTPTSP 337				
RESULT 13					
ID	VRP1 YEAST STANDARD; PRT; 751 AA.				
AC	P37370;				
DT	01-OCT-1994 (REL. 30, CREATED)				
DT	01-OCT-1994 (REL. 30, LAST SEQUENCE UPDATE)				
DT	01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)				
DE	VERPROLIN.				
GN	VRP1.				
OS	SACCHAROMYCES CEREVISIAE (BAKER'S YEAST).				
OC	EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCOMYCETES.				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=A364;				
RM	95058201				
RA	DONNELLY S.F.H., ROCKLINGTON M.J., PALLOTA D., ORR E.J.				
RL	MOL. MICROBIOL. 10:585-596(1993).				
CC	-1- FUNCTION: INVOLVED IN CYTOSKELETAL ORGANISATION AND CELLULAR				
CC	GROWTH. MAY EXERT ITS EFFECTS ON THE CYTOSKELETON DIRECTLY, OR				
CC	INDIRECTLY VIA PROLINE-BINDING PROTEINS (E.G. PROFILIN) OR				
CC	PROTEINS POSSESSING SH3 DOMAINS.				
DR	EMBL, Z26645; SCVERPRL.				
DR	PIR, S39626; S39626.				
KM	REPEAT.				
FT	DOMAIN 5 14 POLY-PRO.				
FT	DOMAIN 239 245 POLY-PRO.				
FT	DOMAIN 349 357 POLY-PRO.				
FT	DOMAIN 396 406 POLY-PRO.				
FT	DOMAIN 424 431 POLY-PRO.				
FT	DOMAIN 462 468 POLY-SER.				
FT	DOMAIN 704 708 POLY-PRO.				
QO	SEQUENCE 751 AA; 75953 MW; 3267787 CN;				

```

Query Match          4.7%; Score 116; DB 8; Length 751;
Best Local Similarity 28.4%; Pred. No. 3,63e-02;
Matches 19; Conservative 23; Mismatches 23; Indels 2; Gaps 2;

Db 374 pcdlapp1nntcvsyppn-kasemppppppppppgafcsaalsaaeip1nql-pppppp 431
      :::: ::1:111 :: 1:1 11 1: ::::: 111 1 1:1
Qy 271 ABDISSTSDTGSPLPMLQPCVSPSPFTHPTGGYTLPLPPTLPTVVOQLHLLPDSAP 330

Db 432 svatsvcp 438
      :: :: 1
Qy 331 TPTPTSP 337

RESULT 14
ID YKE9 YEAST STANDARD; PRT; 441 AA.
AC P35728;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE)
DE HYPOTHETICAL 49.6 KD PROTEIN IN FBI-10A2 INTERGENIC REGION.
GN YKL059C OR YKL316.
OS SACCAROMYCES CEREVISIAE (BAKER'S YEAST).
OC EUKARYOTA; FUNGI; ASCOMYCOTINA; HEMIASCOCYCETES.
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=5288C;
RM 94378723
RA RASMUSSEN S.W.;
RL YEAST 10:563-568 (1994).
DR EMBL; X75781; SCX1286K.
DR EMBL; Z28059; SCYK1055C.
DR PIR; S37881; S37881.
DR PIR; S39179; S39179.
DR PIR; S44524; S44524.
DR PROSITE; PS00995; TCPP_3.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 441 AA; 49647 MW; 1018731 CN;

Query Match          4.6%; Score 115; DB 8; Length 441;
Best Local Similarity 21.2%; Pred. No. 4,76e-02;
Matches 25; Conservative 25; Mismatches 37; Indels 5; Gaps 4;

Db 349 ewetflkqeehgsakdngpcktkklnkdgcaglnntelptsvmggtvppppl 408
      ::1:1: 11 :::: 1:1 :: 1:1:1: 111 :: 1:1 1
Qy 242 QIPCYAIRHIEHLNCTR-GIFPGSRRRTLCADPDISGTSPTGSLPNIQPCYSPPTHP- 299

Db 409 pfqiprfmfpmfmpctatic-nphqadsp 439
      1 1 ::1:1 11: : 1 1:1:
Qy 300 PTC--QYTLPLPTLPTTVQQLHLLPDSA 329

RESULT 15
ID POLR KYMWJ STANDARD; PRT; 1874 AA.
AC P36304;
DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DT 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE RNA REPLICASE POLYPROTEIN (EC 2.7.7.48).
OS KENNEDY YELLOW MOSAIC VIRUS (STRAIN JERVIS BAY) (KYMW).
OC VIRIDAE; SS-RNA NONENVELOPED VIRUSES; TYMOVIRIDAE.
RN (1)
RP SEQUENCE FROM N.A.
RC RM 90218040
RA DING S.; KEESE P.; GIBBS A.;

```

